

# UC Irvine

## UC Irvine Previously Published Works

**Title**

Analysis of gap gene regulation in a 3D organism-scale model of the *Drosophila melanogaster* embryo.

**Permalink**

<https://escholarship.org/uc/item/77n6h8h1>

**Journal**

PloS one, 6(11)

**ISSN**

1932-6203

**Authors**

Hengenius, James B  
Gribskov, Michael  
Rundell, Ann E  
et al.

**Publication Date**

2011

**DOI**

10.1371/journal.pone.0026797

Peer reviewed

# Analysis of Gap Gene Regulation in a 3D Organism-Scale Model of the *Drosophila melanogaster* Embryo

James B. Hengenus<sup>1</sup>, Michael Gribskov<sup>1</sup>, Ann E. Rundell<sup>2</sup>, Charless C. Fowlkes<sup>3</sup>, David M. Umulis<sup>4\*</sup>

<sup>1</sup> Department of Biological Sciences, Purdue University, West Lafayette, Indiana, United States of America, <sup>2</sup> Department of Biomedical Engineering, Purdue University, West Lafayette, Indiana, United States of America, <sup>3</sup> Department of Computer Science, University of California Irvine, Irvine, California, United States of America, <sup>4</sup> Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana, United States of America

## Abstract

The axial bodyplan of *Drosophila melanogaster* is determined during a process called morphogenesis. Shortly after fertilization, maternal *bicoid* mRNA is translated into Bicoid (Bcd). This protein establishes a spatially graded morphogen distribution along the anterior-posterior (AP) axis of the embryo. Bcd initiates AP axis determination by triggering expression of gap genes that subsequently regulate each other's expression to form a precisely controlled spatial distribution of gene products. Reaction-diffusion models of gap gene expression on a 1D domain have previously been used to infer complex genetic regulatory network (GRN) interactions by optimizing model parameters with respect to 1D gap gene expression data. Here we construct a finite element reaction-diffusion model with a realistic 3D geometry fit to full 3D gap gene expression data. Though gap gene products exhibit dorsal-ventral asymmetries, we discover that previously inferred gap GRNs yield qualitatively correct AP distributions on the 3D domain only when DV-symmetric initial conditions are employed. Model patterning loses qualitative agreement with experimental data when we incorporate a realistic DV-asymmetric distribution of Bcd. Further, we find that geometry alone is insufficient to account for DV-asymmetries in the final gap gene distribution. Additional GRN optimization confirms that the 3D model remains sensitive to GRN parameter perturbations. Finally, we find that incorporation of 3D data in simulation and optimization does not constrain the search space or improve optimization results.

**Citation:** Hengenus JB, Gribskov M, Rundell AE, Fowlkes CC, Umulis DM (2011) Analysis of Gap Gene Regulation in a 3D Organism-Scale Model of the *Drosophila melanogaster* Embryo. PLoS ONE 6(11): e26797. doi:10.1371/journal.pone.0026797

**Editor:** Johannes Jaeger, Centre for Genomic Regulation (CRG), Universitat Pompeu Fabra, Spain

**Received:** April 28, 2011; **Accepted:** October 4, 2011; **Published:** November 16, 2011

**Copyright:** © 2011 Hengenus et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was funded by the National Institutes of Health (NIH) Biophysics Training Grant (5T32GM008296-21). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dumulis@purdue.edu

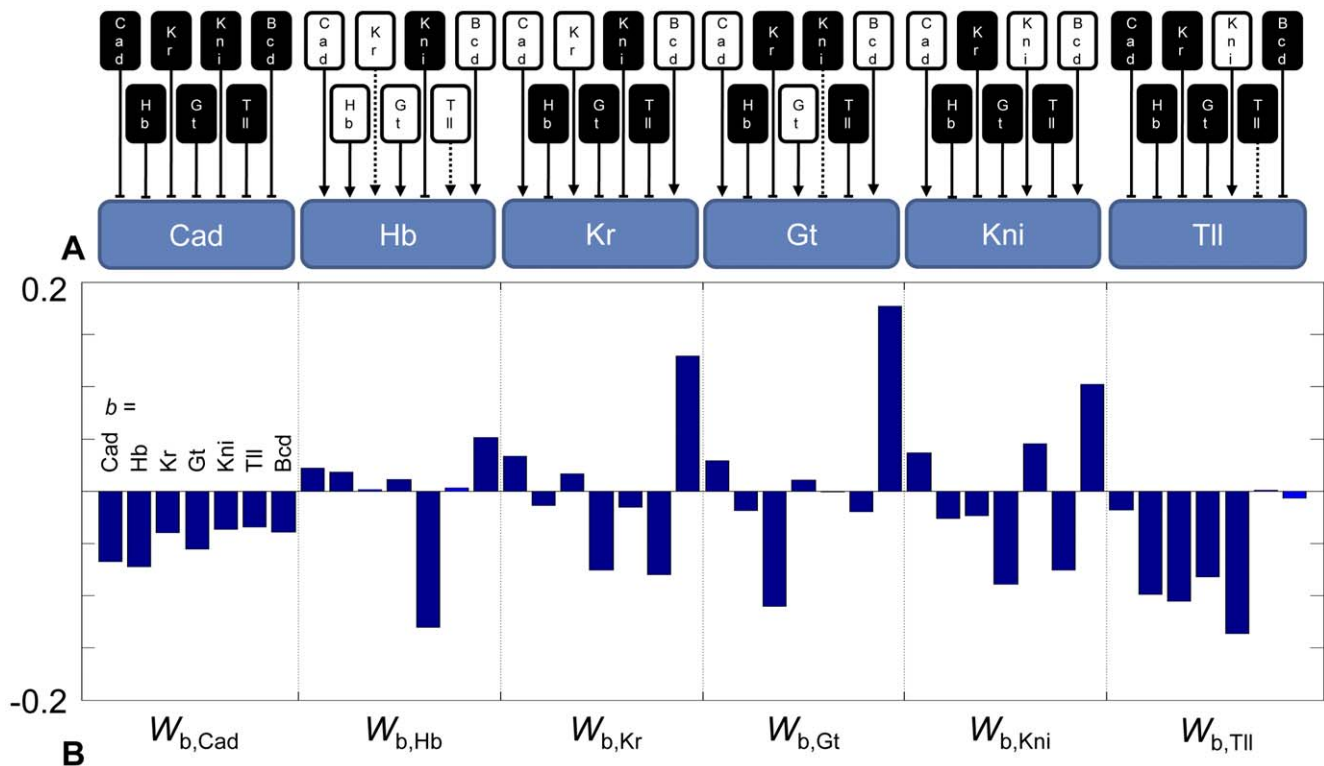
## Introduction

Embryonic development in *Drosophila melanogaster* is initiated with the formation of spatial morphogen distributions in the early embryo. The dynamic spatial patterns of diffusive morphogens encode information which specifies organism-scale development [1,2]. Nonuniform initial spatial distributions of maternally deposited morphogen mRNAs, coupled with diffusion, decay, and complex genetic regulatory interactions, give rise to finer patterns that subdivide the dorsal-ventral (DV) [3–5] and anterior-posterior (AP) axes [2,6] into distinct developmental regions.

The gap gene system is one of the most widely studied morphogen systems in *Drosophila* and is involved in delineation of boundaries of gene expression within the AP body plan [2]. AP patterning events begin approximately one hour post-fertilization. This patterning foreshadows the subsequent segmentation of the embryo [1,2,6–9]. During early development, the embryo is a polynucleated syncytium; most nuclei are arrayed in a thin layer near the surface of the embryo. Due in part to a cytoplasmic viscosity gradient common to insect embryos [10], morphogens (here, gap gene products) are thought to diffuse freely through periplasm near the embryonic surface and less substantially through the interior. Here, they regulate transcription within the periplasmic nuclei [2]. The process is initiated by the gene products of maternally-deposited, spatially-heterogeneous *bicoid*

(Bcd), *caudal* (Cad), and *nanos* mRNAs [2,11,12]. Maternally deposited RNA species regulate expression of the gap genes: *Hunchback* (Hb, with a maternal mRNA contribution), *Giant* (Gt), *Tailless* (Tll), *Krüppel* (Kr), and *Knirps* (Kni) (see **Fig. 1a**) [11,13,14]. The gap genes, in turn, regulate the pair-rule genes which in turn control segment-polarity genes and embryonic segmentation [1,2,6,15].

Most inferences regarding the gap genetic regulatory network (GRN) have been drawn from mutant and gene dosage studies in which the effects on morphology, gap, pair-rule, or segment polarity genes are observed [12,16–36]. While these experiments are informative, it is difficult to unambiguously derive genetic regulatory interactions from such data; phenotypic changes may arise via direct action of the perturbed gene or via downstream targets of that gene. In contrast, Reinitz, Jaeger, and others applied a reverse engineering approach using dynamic wild-type data. Computational studies have modeled gap gene patterning using 1D partial differential equation (PDE) systems or ordinary differential equation systems that include an implicit approximation to the PDE [13,14,37–40] and logical rule sets [41]. These models represent the lateral trunk region of the *Drosophila* embryo along the AP axis, typically omitting the anterior and posterior end regions (with the exception of [40]). GRN topology is represented by a regulatory weight matrix and gene expression is modeled by a transfer function that sums the regulatory impact of each



**Figure 1. Gap gene genetic regulatory network.** The model representation of the gap gene network. The network topology in (A) represents negative (black box, flat line) and positive (white box, arrowhead line) regulatory effects on each target gene (blue). Dashed lines represent near-zero regulatory inputs that may be negligible. This qualitative topology is quantified in (B) as a set of genetic regulatory network (GRN) weight parameters  $w_{b,a}$ , the influence of gene  $b$  on gene  $a$ . From left to right, each set of seven inputs represent Cad, Gt, Hb, Kni, Kr, Tll, and Bcd. Each cluster of seven interactions represents a target gene Cad, Gt, Hb, Kni, Kr, and Tll.  
doi:10.1371/journal.pone.0026797.g001

regulatory protein on expression of the others (see **Fig. 1b**) [13]. Model-driven inferences about GRN topology (i.e., inferring whether and to what degree one morphogen regulates expression of other morphogens) have been obtained by inverse modeling: optimizing the regulatory weight matrix against experimental gap gene expression data in hopes of recovering “true” GRNs [14,42–45]. Findings have been mixed. Biological systems are thought to be robust (and thus insensitive) to perturbations. Some GRN parameters are highly sensitive while considerable uncertainty is associated with others [44,45].

Previous 1D PDE models have been used effectively to infer network topology and investigate patterning regulation [13,14,42, 45,46], but there are some questions that are better investigated using a full 3D spatial patterning model. Many important 3D effects, including variable diffusive path lengths around the embryo surface and optimization against 3D data, cannot be observed in a 1D model domain. DV asymmetries in gap gene distribution and possible interactions between the gap gene system and DV patterning systems are also neglected. Further, these 3D data may serve to constrain GRN optimization and inference.

Quantitative spatiotemporal atlases of gene expression data in the *Drosophila* embryo have been published and provide the starting point for quantitative analysis. [47–49]. The atlas includes measurement of gap gene expression collected from hundreds of individual embryos and registered onto a standardized 3D mesh of nuclei coordinates using pair-rule gene expression patterns as fiduciary points (mesh coordinates available in File S1). This composite VirtualEmbryo (VE) is a logical starting point for the development of 3D embryonic GRN models. It provides a ready-made embryonic

geometry for full spatial PDE representations of the gap gene system. It also contains quantitative expression data against which we can optimize model parameters (and thus infer GRNs).

Using the VE data, we evaluate the impact of 1D model assumptions, conversion from 1D to 3D geometries, and incorporation of fully 3D protein distribution data in model simulation. Herein we reconstruct the 1D gap gene model of Jaeger *et al.* [13] using the finite element method (FEM) and extend it to the 3D VE geometry (**Fig. S1**). The 1D model of Jaeger *et al.* [13],  $M^{1D-P}$  (see Table 1 for model definitions), is refit to lateral expression data from the VE. We then extend the 1D model PDEs to the full 3D embryonic geometry described by Fowlkes *et al.* and compare GRNs inferred from 1D and 3D models. Though 1D models focus on the lateral AP axis in 1D simulations, gap genes are not uniformly distributed along the DV axis. Coupled with the 3D geometry, DV asymmetries in initial conditions may encode positional information partially responsible for the observed AP patterning. As a preliminary exploration of asymmetric DV effects in an embryonic geometry, we evaluate the model using DV-asymmetric Bcd concentration data from thirteen embryos compiled in the VE.

In addition to GRN sensitivities highlighted by previous 1D analyses [14,38,39,44,50], we find that the 3D model exhibits fragility with respect to the shape of maternal gradients: GRNs which were inferred by optimization of 1D models showed similar gap gene patterning when applied to 3D models with DV-symmetric Bcd. However, these GRNs gave rise to qualitatively different patterns in DV-asymmetric models. These realistic Bcd gradient models also captured some of the DV-asymmetries in gap

gene patterning. The 3D models were also sensitive to small perturbations in GRNs; regulatory networks which were qualitatively similar (i.e., all network interactions maintained the same excitatory or inhibitory relationships and differed only by small changes in magnitude) led to qualitatively different gap gene patterns. Refitting of the DV-asymmetric 3D model to VE data produced a GRN which was similar to 1D GRNs but which produced an improved fit.

Another question addressed in this study is whether inclusion of 3D data improves optimization by the inclusion of additional constraints without increasing the degrees of freedom in the model. Unexpectedly, we found that the incorporation of additional 3D information in the form of a realistic DV-asymmetric Bcd worsened the error between optimized 3D models and data. This suggests the involvement of additional regulators in the formation of DV-asymmetries and indicates a direction for future modeling studies.

## Results

### One-Dimensional Model Analysis

Before analyzing the effects of embryonic geometry and DV-asymmetric positional information, we reimplemented the 1D model of Jaeger *et al.* using the finite element method. In this work we denote model variants with  $M$ ; superscripts represent model domains and subscripts signify initial conditions if multiple initial conditions are used. The 1D model of Jaeger *et al.* is called  $M^{1D-P}$  (using a 1D domain representing a partial AP length of 35%–92%; full model nomenclature available in **Table 1**). We verified  $M^{1D-P}$  against Jaeger *et al.*'s simulated output. Whereas the original model limited gene expression to a finite number of discrete nuclear coordinates along the 35–92% region of the embryonic AP axis, the FEM approximates a continuous solution to these equations along this domain. Discrete versus continuous model comparisons by Gursky *et al.* suggest that embryonic patterning is not strongly coupled to nuclear position and that continuous models are comparable to discrete models of gene expression [51]. Our results agree with this finding. FEM simulations produce model output comparable to Jaeger *et al.*'s discrete 1D model (**Fig. 2a**, dashed line, cf. Figure S20 in [13]).

Though  $M^{1D-P}$  recapitulated previous results when simulated in the region from 35–92% on the AP axis, we sought to determine whether moving the boundaries to the embryo ends perturbed gap gene patterning in the trunk region. It is unclear *a priori* how

modification of boundary conditions might impact the model output, because the selection of boundaries at 35% and 92% in earlier work appears to coincide with either maxima or minima of gap gene distributions; at these positions, spatial derivatives are near zero and diffusive flux may be negligible. Using no-flux boundaries at 0% and 100% EL, coupled with the parameters and initial conditions specified in the original model [13], we evaluated  $M^{1D-P}$  and  $M^{1D-F}$  using the GRN parameters  $P^0$  reported by Jaeger *et al.* [13]. Herein, parameter sets are denoted  $P$  and super- and subscripts have model-specific meanings. The simulated patterns from the original 35–92% AP and the 0–100% AP domains are shown in **Figure 2a**'s dashed and solid lines, respectively.

Pronounced shifts in Tll and Kr distributions, coupled with the qualitative change in the anterior Gt distribution, demonstrate the role boundary conditions play in the distribution of gap gene products for a given set of parameters. Though the output of  $M^{1D-P}$  qualitatively resembles the expression data collected previously when evaluated at  $P^0$  [13], these findings suggest that  $M^{1D-P}$ 's agreement with data arises from a combination of the inferred GRN and the domain's boundary conditions. Thus, the internal zero-flux boundary conditions used in previous models may bias GRN inference. To evaluate the impact of boundary placement on GRN inference, we performed a numerical gradient descent search of the parameter space to minimize the root mean squared error (RMSE) between  $M^{1D-F}$  and  $M^{1D-P}$  (represented by the dashed line in **Figure 2a**). The search was initialized with the previously reported optimal  $P^0$ . The result of this search, optimized GRN  $P_{1D-P}^{1D-F}$  (superscript denotes the model being optimized and subscript denotes data with which the model is optimized), is illustrated in **Figure 2b**. Here, the output of  $M^{1D-P}$  represents extant models' with internal no-flux boundaries.

Though domain boundary placement affects the banding pattern, **Figure 2b** suggests that these constraints have a limited effect on GRN inference. Optimizing the GRN parameters of  $M^{1D-F}$  to fit the original model output recovered a quantitatively similar patterning within the 35–92% AP length of the full 1D domain. Additionally, the optimized GRN  $P_{1D-P}^{1D-F}$  was qualitatively similar to  $P^0$  (e.g., though optimized parameters underwent small changes in magnitude, all parameters maintained the same sign, **Fig. 2e**).

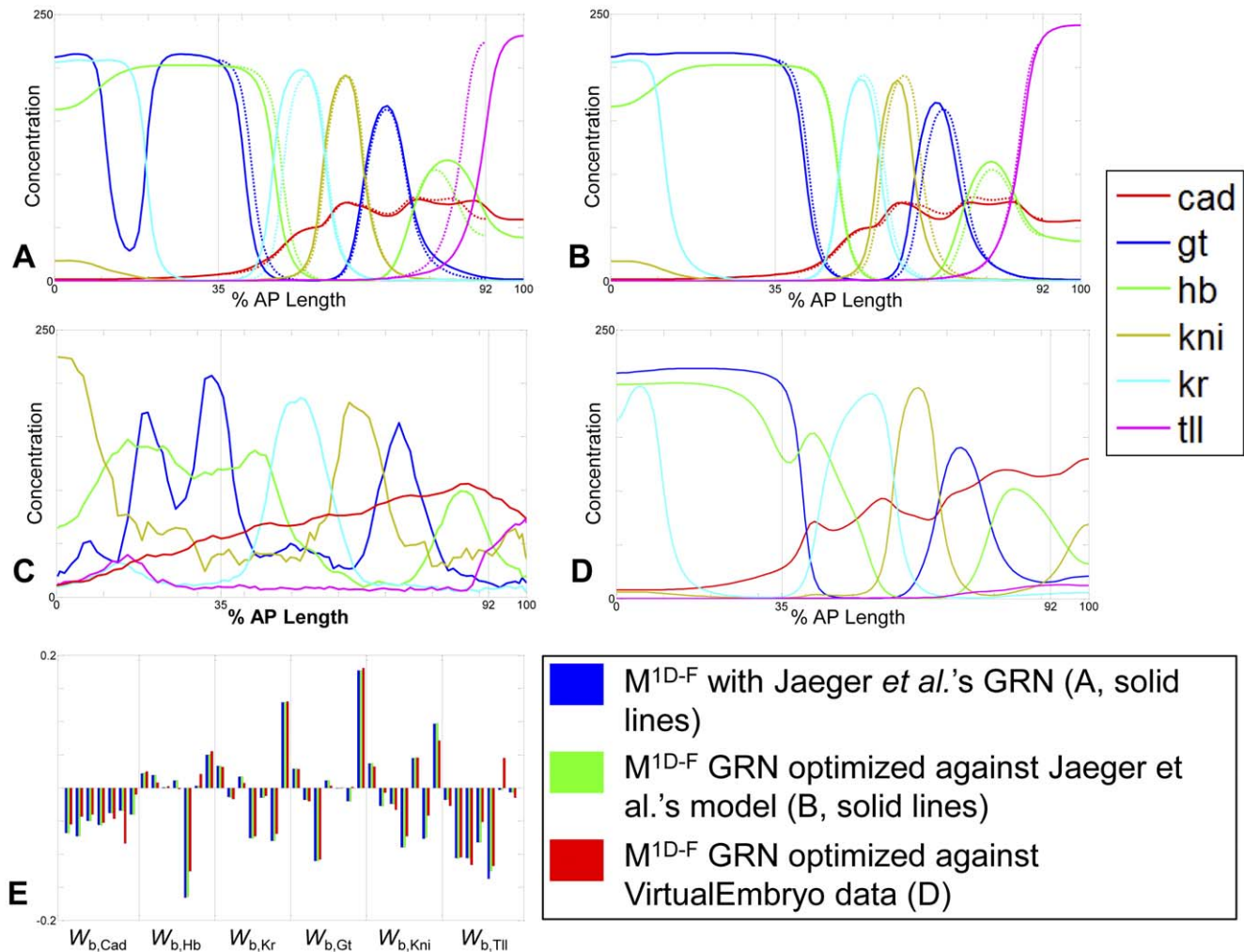
To facilitate a direct comparison between 1D and 3D models presented herein, we first evaluated the goodness-of-fit between the 35–92% AP ( $M^{1D-P}$ ) and full AP domain ( $M^{1D-F}$ ) 1D

**Table 1. Model Variants and Corresponding Optimal Parameter Sets.**

Model	Geometry	Initial Conditions	Optimal GRN Parameters* ( $M_{Data}^{Model}$ )
$M^{1D-P}$	1D domain representing partial 35%–92% AP axis	$Gt_0 = Kni_0 = Kr_0 = Tll_0 = 0$ ; Bcd <sub>SS</sub> , Hb <sub>0</sub> , Cad <sub>0</sub> values in Jaeger <i>et al.</i> [13]	N/A (evaluated with Jaeger <i>et al.</i> 's reported GRN, $P^0$ )
$M^{1D-F}$	1D domain representing full 0%–100% AP axis	$Gt_0 = Kni_0 = Kr_0 = Tll_0 = 0$ ; Bcd <sub>SS</sub> , Hb <sub>0</sub> , Cad <sub>0</sub> values in Jaeger <i>et al.</i>	$P_{1D-P}^{1D-F}$ (fit to Jaeger's model output); $P_{VE}^{1D-F}$ (fit to VirtualEmbryo data)
$M_{Bcd1D}^{3D}$	VE 3D domain	$Gt_0 = Kni_0 = Kr_0 = Tll_0 = 0$ ; Bcd <sub>SS</sub> , Hb <sub>0</sub> , Cad <sub>0</sub> values in Jaeger <i>et al.</i> projected about AP axis (see <b>Fig. 3b–d</b> )	$P_{Bcd1D}^{3D}$ (evaluated with GRN $P_{VE}^{1D-F}$ )
$M_{Bcd3D}^{3D}$	VE 3D domain	$Gt_0 = Kni_0 = Kr_0 = Tll_0 = 0$ ; Hb <sub>0</sub> , Cad <sub>0</sub> values in Jaeger <i>et al.</i> projected about AP axis. Bcd <sub>SS</sub> interpolated from VE data ( <b>Fig. 3e</b> ).	$P_{Bcd3D}^{3D}$ (evaluated with GRN $P_{VE}^{1D-F}$ )
$M_{Bcd3D-S}^{3D}$	VE 3D domain	$Gt_0 = Kni_0 = Kr_0 = Tll_0 = 0$ ; Hb <sub>0</sub> , Cad <sub>0</sub> values in Jaeger <i>et al.</i> projected about AP axis. [Bcd] <sub>SS</sub> interpolated from VE data and smoothed ( <b>Fig. 3f</b> ).	$P_{Bcd3D-S}^{3D}$

\*optimized by fitting model output to Virtual Embryo data unless otherwise noted.

doi:10.1371/journal.pone.0026797.t001



**Figure 2. One-dimensional model results.** Model output was simulated over a 0–100% AP length domain using the optimal GRN reported by Jaeger *et al.* Solid vertical lines represent the original model boundaries, not used in this simulation. (A)  $M^{1D-F}$  (solid lines) shows qualitative agreement with the Jaeger model  $M^{1D-P}$  (dashed lines) in the 35–92% AP range, but shows discrepancies at either end of the domain due to the movement of boundaries; all species displayed at  $t = 70$  min. (B) The best-fit GRN from Jaeger *et al.* was locally optimized to improve the agreement of the 0–100% AP length, model  $M^{1D-F}$  (solid lines), and the original Jaeger *et al.* original model ( $M^{1D-P}$  dashed lines); all species displayed at  $t = 70$  min. (C) VE protein data for Gt, Hb, Kni, Kr at  $t = 70$  min; VE mRNA data for Tll at  $t = 70$  min; protein data from Jaeger *et al.* for Cad at  $t = 56$  min. (D) Model output ( $M^{1D-F}$ ) was also optimized against VE data (RMSE = 13.992); Gt, Hb, Kni, Kr, Tll at  $t = 70$  min; Cad at  $t = 56$  min. Despite modest improvements in model agreement in the 35% and 92% region (C–D), the resulting changes in parameter values were small. (E) Optimized parameter magnitudes vary but signs remained the same in most cases (blue -  $P^0$ ; green -  $P_{1D-P}^{1D-F}$ ; red -  $P_{VE}^{1D-F}$ ). doi:10.1371/journal.pone.0026797.g002

models using VE data. When possible, we use protein expression data from the VE: Gt, Hb, Kni, and Kr protein data is available across six equidistant time points spanning 50 minutes. Tll protein data is unavailable and we use Tll mRNA data as a surrogate for the protein distributions. Cad protein distributions are also unavailable in the VE; we substitute 1D Cad data from Jaeger *et al.* [13] that spans 45 minutes with seven time points. Because the 1D model domains represent the lateral region of the full embryo, we extracted expression data from this region of the VE (Fig. 2c). We performed a constrained search of GRN parameters initialized at  $P_{1D-P}^{1D-F}$  to yield an optimized GRN  $P_{VE}^{1D-F}$  (subscript VE denotes VE training data). The resulting model output and a comparison of model parameters are shown in Figures 2d–e.

Though  $M^{1D-F}$  was capable of recovering the output of  $M^{1D-P}$  (with parameter set  $P_{1D-P}^{1D-F}$ ) and VE data ( $P_{VE}^{1D-F}$ ) within the 35–92% AP axis, poor fit to VE data persisted outside of this region. The 0–35% and 92–100% AP regions exhibit qualitative

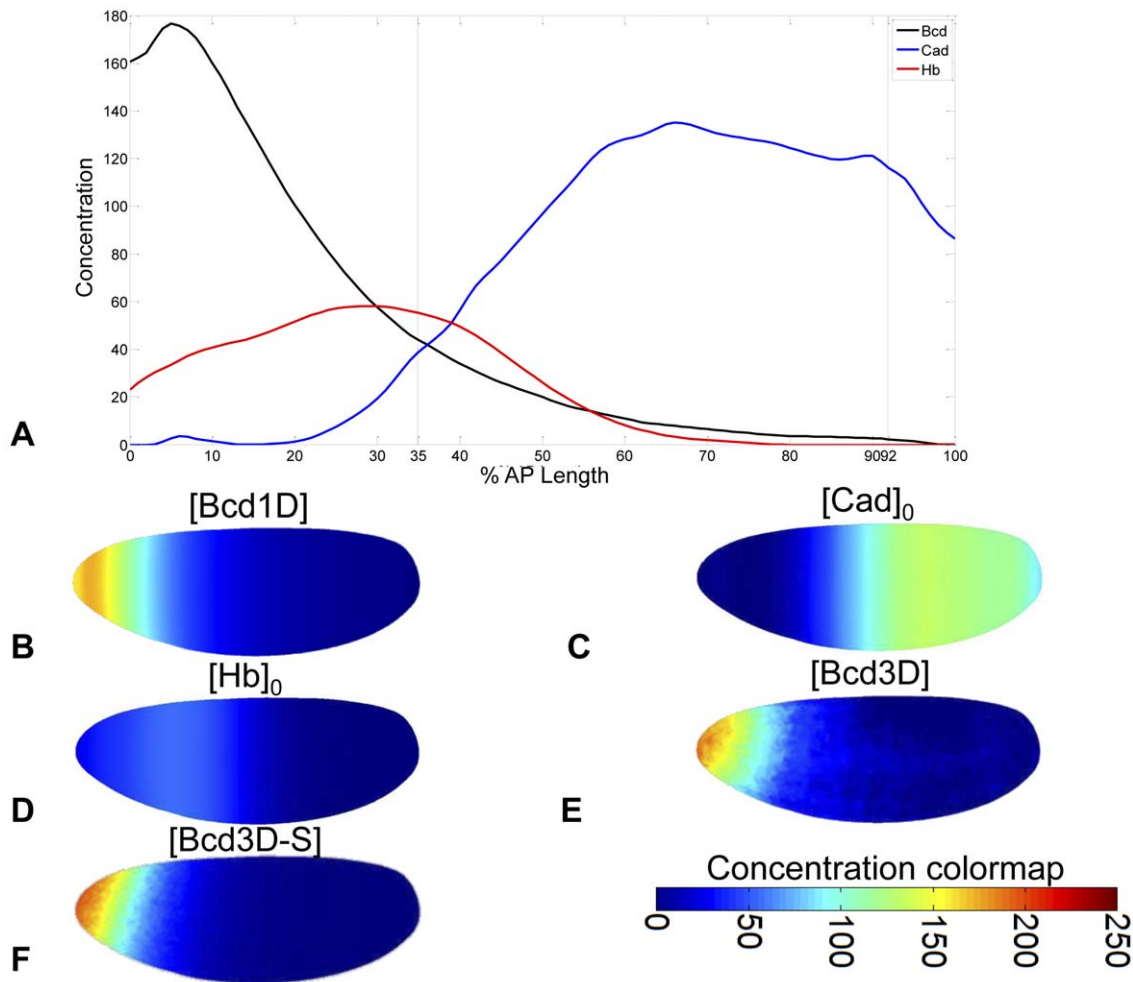
disagreement with VE data in these regions consistent with the biological requirement for additional head and tail patterning genes (Fig. 2c–d).

### Three-Dimensional Model Analysis

Beginning with the GRN optimized on the full 1D domain, we extended the model to a 3D domain using the geometry in the VE. This was performed by implementing the system of PDEs on a 2D surface “wrapped” around the VE geometry. We used this model to evaluate the effects of both model geometry and DV-asymmetric initial conditions on model output.

To assess the effects of model geometry on patterning independent of initial conditions, the model was first simulated using DV-symmetric initial conditions ( $M_{Bcd1D}^{3D}$ ): Bcd, Hb, and Cad distributions at time zero were obtained from the original 1D model and projected around the surface of the embryo (Fig. 3a–d). Evaluated at the previously inferred optimal 1D GRN ( $P_{VE}^{1D-F}$ ),





**Figure 3. 1D and 3D initial conditions.** Initial conditions in various models. (A) 1D model initial conditions, reported by Jaeger *et al.*, and used in models  $M^{1D-P}$  and  $M^{1D-F}$ . (B) 1D initial conditions were mapped onto the 3D embryonic geometry ( $M_{Bcd1D}^{3D}$ ). (C), 1D initial Cad protein distribution, (D) 1D initial Hb protein distribution. Subsequent models incorporated (E) DV-asymmetric interpolated [Bcd] distribution ( $M_{Bcd3D}^{3D}$ ) or (F) smoothed DV-asymmetric interpolated [Bcd] distribution ( $M_{Bcd3D-S}^{3D}$ ).  
doi:10.1371/journal.pone.0026797.g003

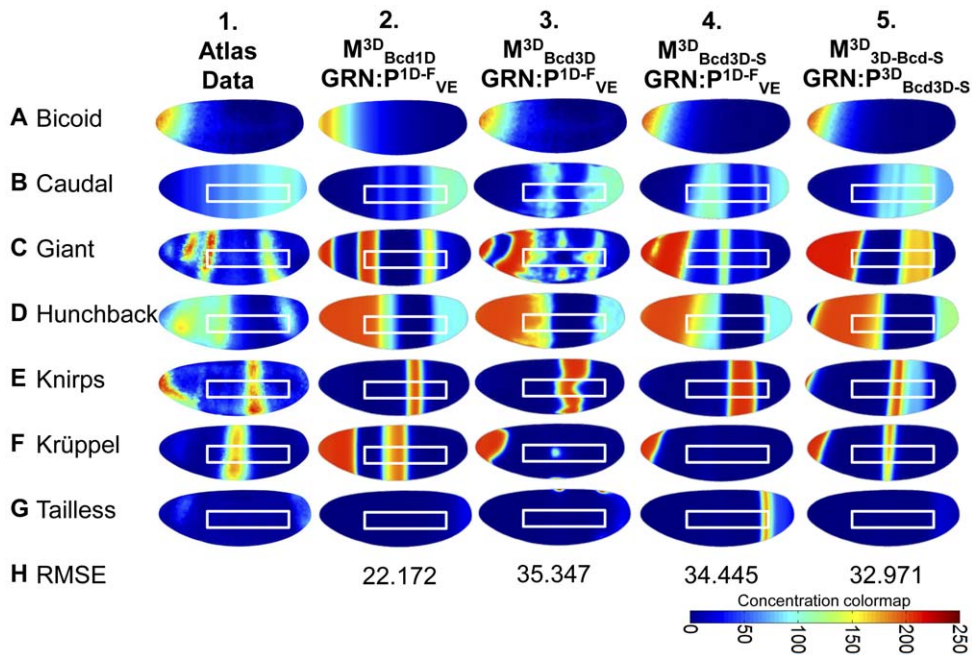
model  $M_{Bcd1D}^{3D}$  yielded patterning qualitatively similar to the full length 1D model output (Fig. 4a–g, column 2). To confirm our derivation of the diffusion constants (see Methods) and rule out unintentional adjustment of the diffusive length constant ( $\sqrt{D_a/\lambda_a}$ ), we performed a continuation of diffusion constants while holding decay parameters ( $\lambda_a$ ) values constant (Figs. S2, S3). While band overlap does vary with diffusion constants, they are quantitatively similar. Interestingly, symmetric Bcd models appear robust against increased diffusion (Fig. S2) while increased diffusion disrupted patterning in asymmetric Bcd models (Fig. S3). The pattern formation timescourse for Bcd-symmetric patterning is animated in Movies S1, S2, S3, S4, S5, S6.

Though there are some DV-asymmetries present in the output (e.g., slight curvature of the anterior Gt stripe), 1D versus 3D domain geometry alone has only a modest impact on DV patterning of gap genes. This suggests that the pronounced DV-asymmetries present in the final distributions of the proteins at the onset of nuclear division 14 (Fig. 4, column 1) stem from other sources. We consider the effect of spatial information encoded in initial DV asymmetries of protein distributions. The coupling of gap gene regulation with DV-patterning systems [5,52,53] is another possibility.

### Effect of Dorsal-Ventral Asymmetric Bcd

To evaluate the impact of DV-asymmetric inputs on the model, we modified the steady-state Bcd distribution shown in Figure 3b to incorporate a realistic DV gradient (Fig. 3e). Unlike other morphogens, the Bcd distribution is static over the entire time course of model simulation. This allowed us to create a single interpolant of VE Bcd data and use it as a model input for all 70 minutes of the simulation. The pattern formation timescourse for Bcd-asymmetric patterning is animated in Movies S7, S8, S9, S10, S11, S12.

Evaluated at the optimal 1D GRN  $P_{VE}^{1D-F}$ , model  $M_{Bcd1D}^{3D}$  produces patterning that is radically different from DV-symmetric 1D ( $M^{1D-F}$ ) and 3D ( $M_{Bcd1D}^{3D}$ ) models (Figs. 4a–g, column 3). The most striking example of this is the Kr model output; whereas Kr forms a full band *in vivo*,  $M_{Bcd3D}^{3D}$  lacks full lateral expression of Kr and has an anomalous region of expression at the anterior end of the embryo (Fig. 4f, column 3). Similarly, the simulated Hb concentrations remain above observed levels (Fig. 4d, column 3). The posterior Hb band also shifts to the posterior end of the embryo. Gt exhibits qualitative disagreement with the VE data; whereas anterior Gt expression is observed only in a limited dorsal region of the embryo (Fig. 4c, column 1), the anterior of the  $M_{Bcd3D}^{3D}$  is saturated with Gt (Fig. 4c, column 3). Further, though



**Figure 4. Three-dimensional model results.** Simulation results in the 3D model. (A–H) Lateral view of VE geometry is shown in rows A–G (Gt, Hb, Kni, Kr, Tll at  $t = 70$  min, Cad at  $t = 56$  min); row H displays RMSE difference between model and VE data summed with all time points. Column 1 shows scaled VE data. Column 2 displays output from  $M^{3D}_{Bcd1D}$  evaluated with GRN  $P^{1D-F}_{VE}$ . Column 3 contains output from  $M^{3D}_{Bcd3D}$  incorporating DV-asymmetric Bcd data and GRN  $P^{1D-F}_{VE}$ ; Column 4 illustrates the effect of the smoothed Bcd interpolant in  $M^{3D}_{Bcd3D-S}$  while considering the same GRN  $P^{1D-F}_{VE}$ . Column 5 displays output from  $M^{3D}_{Bcd3D-S}$  with reoptimized parameters  $P^{3D}_{Bcd3D-S}$ . White boxes indicate the lateral areas where Jaeger *et al.* optimized their 1D model. Animations of pattern development are available for column 2 ( $M^{3D}_{Bcd1D}$ , **Movies S1, S2, S3, S4 S5, S6**) and column 5 ( $M^{3D}_{Bcd3D-S}$ , **Movies S7, S8, S9, S10, S11, S12**) in the supplementary material. doi:10.1371/journal.pone.0026797.g004

the experimentally observed posterior Gt band (**Fig. 4c, column 1**) is predicted by simulation, it exhibits unusual differences in width along the DV axis. As in previous versions of the model, the best agreement between model and data was found in the lateral 35–92% AP region (**Fig. 4b–g, column 3 white boxes**).

The cell-to-cell variability in patterning found for many simulated proteins (e.g., Gt, Cad, and Kni) in  $M^{3D}_{Bcd3D}$  led us to consider the effect of noise in the VE Bcd distribution. Diffusion of Bcd may serve to smooth this variation *in vivo*; our use of a single static Bcd interpolant in  $M^{3D}_{Bcd3D}$  leads to an artificial persistence of the noise found in VE data (**Fig. 3e**). To test for and remove this artificial condition, we created a regularized version of the Bcd interpolant (**Fig. 3f**). This was constructed by building a simple source diffusion decay (SDD) reaction-diffusion model of Bcd alone [18]. This SDD model was fit to VE data and the steady-state solution was used as the smoothed Bcd interpolant. The model incorporating regularized Bcd,  $M^{3D}_{Bcd3D-S}$ , did not show significant improvement over  $M^{3D}_{Bcd3D}$  when evaluated with  $P^{1D-F}_{VE}$  (**Fig. 4a–g, column 4**). However, it did eliminate the cell-to-cell variability present in  $M^{3D}_{Bcd3D}$ . The model's artificial sensitivity to Bcd noise was especially evident in Gt (**Fig. 4c, columns 3–4**). Two anterior and one posterior Gt bands in  $M^{3D}_{Bcd3D}$  changed width and AP position after smoothing of Bcd. This result suggests that while diffusion may serve as a buffer against transient stochastic variations in protein expression and local concentration (in agreement with stochastic simulation [54]), sustained cell-to-cell variability has the potential to disrupt patterning.

Having observed that a GRN inferred on the 1D domain (and lacking DV asymmetries) produces a qualitatively incorrect fit compared to 3D data, we attempted to optimize the GRN with Matlab's constrained search function `fmincon()` initialized at

$P^{1D-F}_{VE}$  (previously used to estimate  $P^{1D-F}_{1D-P}$  and  $P^{1D-F}_{VE}$ ). This approach failed to reduce model error. Fomekong-Nanfack *et al.* demonstrated that 1D gap gene systems are amenable to optimization by evolutionary algorithms [45]. We therefore employed a genetic algorithm (GA) to more broadly survey the parameter space. Do to computational cost, we used a small population size of 20 genomes to search the GRN parameter space (42 parameters), the GA identified an optimal GRN for  $M^{3D}_{Bcd3D-S}$ . The resulting GRN,  $P^{3D}_{Bcd3D-S}$ , led to a reduction in model error and a modest qualitative improvement with respect to 3D data (**Fig. 4, column 5**). The lateral Kr band missing from the 1D-inferred GRN  $P^{1D-F}_{VE}$  (**Fig. 4f, columns 3–4**) is restored (**Fig. 4f, column 5**), though it is not as wide as the experimentally observed band. Tll no longer shows relative over-expression at the posterior end of the embryo (**Fig. 4g, column 5**). Hb continues to exhibit relative over-expression at the anterior end of the embryo, though its posterior band is shifted closer to its correct position (**Fig. 4d, column 5**). Similarly, the anterior distribution of Gt extends beyond the dorsal region observed in the VE (**Fig. 4c, column 1**). However, its posterior band is now located correctly in **Figure 4c, column 5** (though it is wider than the observed protein band). Beyond differences in concentration of individual proteins, DV-asymmetric Bcd causes a notable qualitative difference in the AP position and emergence of protein bands. Compared to the  $M^{3D}_{Bcd1D}$  (**Fig. 4, column 2**), the DV-asymmetric GRNs (**Fig. 4, columns 4–5**) exhibit DV-asymmetries in their output. For example, the dorsal terminus of the anterior Gt band is posterior to its ventral terminus; it is splayed toward the anterior. This behavior agrees with observed data in the anterior half of the embryo, but the expected DV curvature is either absent (posterior Hb **Fig. 4d, column 5**) or inverted in the posterior half of the embryo. For example, Kni,

whose dorsal terminus should exhibit posterior-splaying (**Fig. 4e, column 5**), is inverted. This DV curvature corresponds in direction to the DV asymmetry of Bcd. The absence of reversed splaying in the output in the posterior portion of the model (though present in the data) suggests that the model may be lacking additional posterior determinant(s) affecting the gap gene system.

In the 3D regime,  $M_{Bcd3D-S}^{3D}$  demonstrated considerable sensitivity to small changes in GRN parameter values. The model was simulated after adding normally distributed noise scaled by each parameter value,  $p_i$ , across a range of magnitudes (sample model output in **Fig. 5**). The model gives output qualitatively similar to the optimal GRN  $P_{Bcd3D-S}^{3D}$  only when parameter noise is low (e.g., 0.1%  $p_i$  in **Fig. 5, column 1**). All other simulations, with noise terms of 1%  $p_i$  and higher, yielded drastically and qualitatively different outputs.

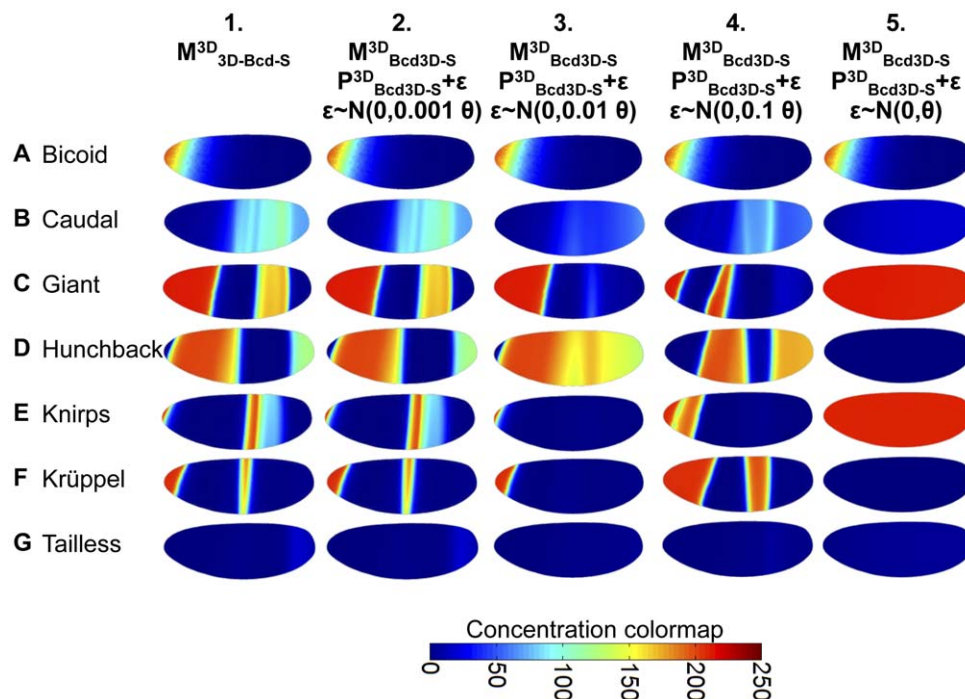
>In summary, the GRNs we inferred in this study are qualitatively similar: magnitudes of parameters vary by approximately 10% and parameter sign stays the same in all but a few low-magnitude parameters (see **Table S1**). A notable exception is the regulatory parameter for the *Kni*→*Tll* interaction; here the sign of the parameter (and thus the regulatory relationship) is reversed. However, we acknowledge that the treatment of *Tll* as a state variable under gap gene regulation is artificial and this biological relevance of this observation is questionable. Optimization leaves most regulatory parameters with the same sign and changes only the magnitudes, and those regulatory weights which change sign have small magnitudes (i.e., small regulatory effects). The use of a global search method (GA) to optimize  $M_{Bcd3D-S}^{3D}$  did not recover a superior GRN that differed qualitatively from the original  $P^0$ .

## Discussion

The understanding of *Drosophila* developmental gene regulation has benefited from advances in quantitative modeling of gene

regulation. However, existing PDE models of AP patterning have been limited to 1D approximations of the 3D geometry. By extending a model of gap gene regulation to a 3D embryonic geometry and adding realistic DV-asymmetry to upstream maternal Bcd, this work allows us to pose new questions about the effects of embryonic shape and DV gradients on gap gene patterning. Jaeger *et al.*'s 2004 model has been succeeded by more recent models of gap gene development incorporating additional regulatory inputs [37–39,46,55–58]. However, recent models of AP patterning retain partial domains (e.g., 35%–92% AP) with internal no-flux boundary conditions and use regulatory schema similar to **eqns 1–3** (see **Methods**) to represent GRNs. We chose the Jaeger *et al.*'s 2004 model as a case study in 1D vs. 3D modeling because it is the representative of many existing 1D models.

Before comparing 1D and 3D geometries, we examined the effect of boundary position in PDE solutions. Though embryos do not contain physical barriers to diffusion at 35% and 92% of the AP axis, small spatial gradients (**Fig. 2a, dashed lines**) at those positions suggested that small diffusive flux would minimize the effects of these internal boundaries. However, we found that the system was sensitive to boundary placement (cf. **Fig. 2a, solid lines**). Though this finding indicates the importance of using biologically realistic boundary conditions (i.e., no-flux boundaries at 0% and 100% AP), the simulations in **Figure 2** also illustrate our limited representation of regulation beyond the 35%–92% trunk region: Omission of terminal gap genes and regulators result in optimized parameter sets that cannot recapitulate expression patterns from 0%–35% and 92%–100% AP in  $M^{1D-F}$  (**Fig. 2a,c**). Optimization to correct the boundary artifacts ( $M^{1D-F}$  with  $P_{1D-p}^{1D-F}$ ) likewise fail to improve agreement with data outside of the 35%–92% region (**Fig. 2b**). The inclusion of terminal gap genes such as Hucklebein in 1D gap gene models [37]



**Figure 5. Model  $M_{Bcd3D-S}^{3D}$  is not robust to noise in GRN parameters.** Parametric noise alters model output. Lateral view of VE geometry for all genes is shown in rows **A–G** (all outputs at  $t = 70$  min). Each column displays  $M_{Bcd3D-S}^{3D}$  output at  $t = 70$  min evaluated with GRN  $P_{Bcd3D-S}^{3D}$ . Columns 2–5 represent randomly chosen sample output when a normally distributed noise vector  $\epsilon$  is added to the GRN parameter set (denoted  $\theta$ ).  $\epsilon$  has mean of 0 and variance that scales with  $\theta$ . doi:10.1371/journal.pone.0026797.g005



provides a basis for extension to full 100% AP 1D and 3D models, though inclusion of Huckebein in a recent 3D modeling study yielded only modest improvements in overall cost and qualitative agreement at the AP extrema [59].

Prior analyses demonstrated the sensitivity of gap gene models to GRN parameter values [14,43,44] and examination of boundary conditions support this finding: GRN parameter optimization corrected boundary artifacts with extremely small changes to parameter values (**Fig. 2e**). Optimization against VE data produced similar small changes in GRN parameters (**Fig. 2e**). The GRN sensitivity of 1D models  $M^{1D-P}$  and  $M^{1D-F}$  was also found in 3D models. **Table S1** collects all parameter values and reports the standard deviation for each parameter across 1D and 3D model optimizations. Parameter  $w_{Gt,Bcd}$  exhibits the highest deviation across models with a standard deviation of 0.05, but this represents only 13% of the total parameter range ( $[-0.2,0.2]$ ). These small changes in GRN parameters do more than shift protein band location as observed in **Figure 2**; they are capable of effecting qualitative patterning changes (e.g., changing the number of protein bands present on the embryo). For example, the transition from  $P_{VE}^{1D-F}$  to  $P_{Bcd3D-S}^{3D}$  in model  $M_{Bcd3D-S}^{3D}$  leads to the loss of a posterior Gt band and the creation of a posterior Kr band (**Fig. 4c,f, columns 4–5**). **Figure 5** shows randomly selected sample model outputs at  $t = 70$  min with increasing levels of normally distributed noise added to the GRN parameter vector. One percent noise was sufficient to induce qualitatively different banding patterns on the 3D geometry. The qualitative changes in patterning for all but the smallest levels of noise confirm the observations of parameter sensitivity in 1D and 3D models. The extreme sensitivity of model outputs to small changes in GRN parameters challenges analyses of GRN evolution positing phenotypically robust fitness landscapes [60–62]. Unfortunately, the computational expense of PDE models prevented an exhaustive exploration of the GRN parameter space and corresponding approximation of a fitness landscape. The fragility of the gap gene system to GRN perturbations bears further study, especially in its contrast to prevailing thoughts that evolution occurs on networks with highly-connected neutral (selectively equivalent) genotypes.

In addition to the parameter sensitivity and boundary conditions, our work also demonstrate the use of accurate 3D geometry and its effects on model predictions. We found that geometry alone has a limited effect on gap gene patterning: Excepting slight DV-asymmetry brought about by the curvature of the 3D embryo, 1D output from  $M^{1D-F}$  (**Fig. 2d**) and 3D output from  $M_{Bcd1D}^{3D}$  (**Fig. 4, column 2**) display qualitatively similar band position along the AP axis. The path length from anterior to posterior extrema differs with DV position: For example, the distance from anterior to posterior extrema is shorter along the dorsal surface than the ventral surface. We thought that this difference in diffusion distance might account for the anterior splaying displayed in VE data (**Fig. 4, column 1**), but this was not the case.

Though the 3D embryonic geometry was insufficient to explain DV-asymmetries in gap gene data, it allowed us to explore the effect of DV-asymmetric protein distributions on patterning. Notably, the 1D Bcd distribution of  $M_{Bcd1D}^{3D}$  (**Fig. 3b**) differed from the typical dorsal-anterior distribution [63,64] also found in the VE (**Fig. 3e**). Experimental noise in this data led to aberrant patterning in most gap genes in  $M_{Bcd3D}^{3D}$  (**Fig. 4, column 3**), but a regularized version of the distribution (**Fig. 3f**) produced cleaner (though qualitatively incorrect) band appearance and position in  $M_{Bcd3D-S}^{3D}$  (**Fig. 4, column 4**). It also produced anterior-splaying in the anterior bands of Gt, Hb, Kni, and Kr. As previously noted, optimization of the sensitive GRN parameters improved qualita-

tive agreement in model patterning with only small changes to parameter values (**Table S1**).

When considering 3D models and the data associated with them, we endeavored to identify any constraints on model optimization. This model has many degrees of freedom and additional information encoded in the DV asymmetries of gap genes might better guide parameter searches toward accurate GRNs. However, we observed no improvement in RMSE values and failed to find any novel GRNs for DV-asymmetric models.

Though our ensemble of models has led to interesting findings, we acknowledge model limitations. Recent modeling studies recognize that Cad and Tll patterning cannot be completely accounted for by gap genes in existing models; maternal mRNA complicates Cad expression and Tll is under the regulation of additional proteins [38]. Instead, newer models use data interpolants to represent these proteins [38]. The absence of these interpolants in our models may contribute to the unrealistic sensitivity of the 3D model parameters and DV-information. 3D interpolating functions incorporating VE data for Cad and Tll are under development; we will use these to explore the behavior of more recent 1D models on the 3D embryonic geometry.

The primary focus of this work is the comparison of 1D and 3D model geometries. **Figures 2d** and **4, column 2** reveal that differences in model geometry can be accommodated by relatively minor adjustments to GRN parameters. The 3D implementation ( $M_{Bcd1D}^{3D}$ ) exhibits minor DV-asymmetries but otherwise mirrors  $M^{1D-F}$ . However, consideration of AP patterning in three dimensions allows us to address the experimentally observed DV-asymmetry in maternal Bcd and downstream AP morphogens. The inclusion of a DV-asymmetric Bcd signal led to qualitatively different patterning with  $P_{VE}^{1D-F}$  (**Fig. 4, columns 2,4**). This suggests that the assumption of DV and AP independence in previous modeling studies is violated. Parameter sensitivity remained high; parameter optimization made small changes to parameter values but led to significantly improved RMSE error (**Fig. 4, columns 4,5**).

Finally, two cases of DV model mismatch suggest modifications that could be incorporated into future models. First, anterior Gt is more highly expressed on the dorsal side of the embryo *in vivo*, but posterior Gt displays posterior-splaying. This expression localization is not accounted for by Bcd distribution alone and should be addressed in future models that also include input from the DV patterning system downstream of the active Dorsal protein distribution [65,66]. Second, many protein species display DV-asymmetry in terms of anterior or posterior splaying. E.g., Cad bands anterior to the AP midline are anterior-splayed (**Fig. 4b, column 1**) while bands posterior to the AP midline are posterior-splayed. This pattern is observed for all modeled proteins (**Fig. 4, column 1**), though it is lacking in DV-symmetric  $M_{Bcd1D}^{3D}$  (**Fig. 4, column 2**). Addition of DV-symmetric Bcd ( $M_{Bcd1D}^{3D}$ ) restores anterior-splaying aligned with the DV Bcd gradient (**Fig. 4, column 5**). This suggests that a missing posterior determinant may be responsible for posterior-splaying. The posterior maternal morphogen Nanos is a candidate that has not been included in previous models. With interpolated Cad and Tll, future models will explore the effects of posterior determinants such as Nanos [67] and, as examined in prior 1D models, Huckebein [37].

## Methods

### Model Construction

Building on the successful 1D/3D embryonic modeling approach of Umulis *et al.*, [4,68], we reimplemented the Jaeger *et al.* model of gap gene regulation ( $M^{1D-P}$ ) using the finite element

method (FEM). This model represents six gene products as state variables: Cad, Gt, Hb, Kr, Kni, and Tll [13]. A seventh protein, Bcd (Bcd), is maintained at a constant concentration during gap gene patterning and is represented as a spatially heterogeneous stationary input [13,63]. Each of the state variables is represented by a PDE,

$$\frac{\partial c_a}{\partial t} = D_a \nabla^2 c_a + R_a \phi_a(u_a) - \lambda_a c_a, \quad (1)$$

where  $c_a$  is the concentration of protein  $a$ , the first term on the right hand side represents diffusion, the second term represents gene expression, and the third term represents first order decay [13].  $D_a$  is the diffusion constant of protein  $a$  and  $\lambda_a$  is the first order decay constant of protein  $a$ .  $R_a$  is the maximal rate of gene expression of proteins  $a$  and  $\Phi_a$  is a sigmoid function,

$$\phi_a(u_a) = \frac{1}{2} \left( \frac{u_a}{\sqrt{u_a^2 + 1}} + 1 \right), \quad (2)$$

which ranges from zero to one and accepts a regulatory argument  $u_a$ :

$$u_a = h_a + \sum_b w_{b,a} c_b. \quad (3)$$

Here,  $h_a$  is a minimal regulatory threshold for expression,  $w_{b,a}$  is an element in the regulatory matrix  $W$  representing the influence of protein  $b$  on the expression of protein  $a$  (ranging from  $-0.2$  to  $0.2$ ), and  $c_b$  is the local concentration of protein  $b$ . There are six PDEs representing protein proteins  $a = \text{Cad, Gt, Hb, Kr, Kni, Tll}$  (eqn. 1). In each PDE, the regulatory effects of all seven proteins,  $b = [\text{41 Kr, Kni, Tll, Bcd}]$ , control protein expression (eqns. 2–3). PDEs are numerically solved using the FEM implemented in the software package COMSOL Multiphysics 3.5a [69]. Except for GRN parameters  $w_{b,a}$ , these parameters are fixed at values in Jaeger *et al.* [13] and may be found in **Table S2**.

Note that previous 1D models were simulated by the spatially-discretized ordinary differential equations using the finite difference method: concentrations were tracked at uniformly-distributed nodes (nuclei) along the AP axis and diffusive fluxes across the  $\Delta x$  inter-node distance were modeled as a first-order differential equations. As such, previously reported diffusion parameters ( $\dot{D}_a$ ) were in units of inverse time [ $1/t$ ]. To convert these parameters to diffusion constants ( $D_a$ ) with units of squared-length-per-time [ $L^2/t$ ], we multiplied  $\dot{D}_a$  by  $(\Delta x)^2$ . To compute  $\Delta x$ , we took into account the length of the original model's domain (0.57 EL) and the number of nodes where the finite difference model was solved (58 nuclei). From these values, we approximated  $\Delta x$  as 0.57EL/57. The model spans 0.35–0.92 or 0.57 EL and is divided into 57 intervals between 58 nodes. In the case of the 3D geometry, we further accounted for the curvature of the embryo in our approximation of  $\Delta x$ . Scaling the embryo length to unity (1 EL), we observed an arc length of 1.14 along the lateral AP. Upon the assumption that curvature was uniformly-distributed along the AP axis,  $\Delta x$  was computed as (0.57/1.14)EL/57. The approach slightly overestimates  $D_a$  in the 3D model relative to 1D because most curvature occurs at the AP extrema and not the trunk, but this does not translate to a large impact on AP patterning versus 1D. Whereas finite difference models explicitly modify  $\dot{D}_a$  values to account for mitotic nuclear division and the halving of  $\Delta x$ , the continuous FEM representation renders diffusion constants independent of nuclear density. It should be noted that this repre-

sentation does not account for reduced effective diffusivity due to increased nuclear trapping. While nuclear density has been linked with decreased effective diffusivity in some simulations of Bcd diffusion [70], Grimm and Wieschaus found that transcription factor distributions are largely independent of nuclear density [71]. 3D nuclear density distributions have been published [47] and nuclear density-dependent diffusion is an area for further investigation.

We developed two FEM meshes on which to simulate spatio-temporal gap gene evolution. A 1D linear domain represents the 35–92% AP axis, and replicates the domain used in previous models [13]. By scaling diffusion constants and choosing initial conditions, the 1D domain also represents the 0–100% AP length ( $M^{1D-F}$ ). A 3D mesh modified from the VE geometry represents a realistic embryonic geometry. Though the embryonic syncytium includes the yolk interior of the embryo, nuclei are located within the periplasmic domain of the exterior surface [10,49]. Cytoplasmic viscosity increases in the embryonic interior and is presumed to limit effective diffusion of gap gene products to the 2D layer in the periplasmic volume containing the nuclei. While some gap gene products may diffuse into yolk, this process may be considered as part of the decay terms,  $\lambda_a$ . We took this into account when constructing the 3D domain. The reaction-diffusion equations (eqns. 1–3) are implemented as weak form PDEs on a 2D manifold (**Fig. S1**); this manifold is “wrapped” around the 3D embryonic geometry in 3D model implementations ( $M_{Bcd1D}^{3D}, M_{Bcd3D}^{3D}, M_{Bcd3D-S}^{3D}$ ).

Though the 3D domain is a closed surface without AP flux boundaries, the partial ( $M^{1D-P}$ ) and full ( $M^{1D-F}$ ) 1D domains are bounded at both termini by zero-flux conditions. These internal boundaries are unrealistic in the case of the partial AP length domain as there are no such physical barriers in the embryo; they were introduced in previous gap gene models to help account for artifacts in previously inferred GRNs [14,42–44]. In full length 1D models the anterior and posterior ends of the embryo are realistically represented by zero-flux boundaries.

Numerical integration of PDEs requires specification of initial conditions as well as boundary conditions. For purposes of model comparison, we chose initial conditions specified in previous models [13]. On both 1D and 3D domains, the proteins Gt, Kni, Kr, and Tll have initial uniform concentrations of zero. Jaeger *et al.* provide initial non-uniform 1D distributions for Cad and Hb (**Fig. 3a**) [13]. These distributions span the entire AP length and provide initial conditions for both the partial and full length domains. Jaeger *et al.* also provide a constant exponential 1D Bcd distribution for the full AP length. These 1D distributions were used as initial conditions in the 1D models ( $M^{1D-F}$  and  $M^{1D-P}$ ). They were projected onto the 3D domain to approximate full 3D initial conditions ( $M_{Bcd1D}^{3D}$ , **Fig. 3b–d**). This projection was performed using built-in interpolation tools in the Comsol package. Provided AP-coordinates and corresponding concentration values, Comsol created a linear interpolant of DV-symmetric concentration values along the AP-axis of the 3D geometry.

While the Bcd data provided by Jaeger *et al.* describes the lateral AP distribution of Bcd, it fails to capture the observed DV asymmetry found in embryonic Bcd. Though sufficient for a 1D model (**Fig. 3a**), the resulting 3D distribution (**Fig. 3b**) qualitatively disagrees with VE data (**Fig. 3e**). We therefore built an interpolating function from the VE Bcd data and used this interpolant when simulating the model ( $M_{Bcd3D}^{3D}$ ). Again, we used Comsol's interpolation functionality. However, this interpolant required full 3D specification of coordinates. We used the coordinates of nuclei and corresponding Bcd concentration values provided in the VE. Because the software does not support interpolation on a 2D boundary (the periplasmic space) in a 3D

geometry, we used nearest-neighbor interpolation (**Fig. 3e**). Because this Bcd distribution is represented in the model as a static interpolant, noise in the data (and hence the interpolant) is not smoothed by diffusion and decay. Initial attempts at directly importing VE Bcd data resulted in persistent asymmetries and mottled distributions inconsistent with data (**Fig. 4, column 3**). In an ideal situation, inter-embryo variability would be averaged out of VE data. However, the data set was generated with few replicates (13 embryos for Bcd [49]) and spatial noise remained. To remove this noise from the interpolant, we first fit a steady-state source-diffusion-decay (SDD) model of Bcd production [18] to VE Bcd data on the 3D domain (**Fig. 4a, column 1**). Once we had obtained agreement between this regularized Bcd distribution and the data, we used the solution of the SDD model to create a new interpolant. This smoothed interpolant shown in **Figure 3f** and  $M_{Bcd3D-S}^{3D}$ 's output (**Fig. 4b–g, columns 4–5**) compares favorably with the results  $M_{Bcd3D}^{3D}$  (**Fig. 4b–g, columns 3**).

Spatiotemporal regulation of gap gene expression spans the mitotic nuclear division between nuclear cycle 13 and 14a. For purposes of comparison, we chose to simulate the same time-course as previous models. We begin by simulating the conclusion of cycle 13 for sixteen minutes, mitosis for five minutes, and continue to simulate cycle 14a for the remaining forty-nine minutes [13]. The reaction-diffusion equations (eqn. 1–3) describe the model during interphase. During mitosis, gene expression (the second term in eqn. 1) is set to zero. Molecules may diffuse and decay, but they are not transcribed or translated while the chromatin is compacted for mitotic division.

This set of initial and boundary conditions, coupled with the reaction-diffusion equations and a geometric domain, constitutes a numerically soluble model. To calculate model error, we used a straightforward root mean squared error cost function:

$$J_{RMSE} = \sum_{t=1}^6 \sum_{a=1}^6 \sum_{i=1}^n \sqrt{\frac{(c_{a,\text{mod}}(\theta, i, t) - c_{a,\text{exp}}(i, t))^2}{6n}}. \quad (5)$$

Here,  $\theta$  is the GRN parameter set,  $n$  is the number of data points in the 35%–92% EL region of the embryo,  $a$  is the index of protein species,  $i$  is the index of  $n$  nuclear coordinates, and  $t$  is the time index. This function sums the root squared error between model output from a given GRN,  $c_{a,\text{mod}}(\theta, i, t)$ , and experimental data,  $c_{a,\text{exp}}(i, t)$ , over data points  $i$ , model proteins  $a$ , and time  $t$ .

$M^{1D-P}$  was originally fit to immunofluorescence data in Jaeger *et al.* [13]. As a result, both the model's concentration units and GRN parameters are scaled to reflect observed relative intensity ranges of those data. To facilitate fitting between models utilizing Jaeger *et al.*'s parameters and VE data, we pre-scaled the VE data to agree with the initial conditions reported by Jaeger *et al.* This was performed by optimizing scaling factors  $A_a$  and offsets  $b_a$  such that the difference between Jaeger *et al.*'s initial conditions and the VE data was minimized,

$$\min_{A,b} \sqrt{((A_a c_{VE,a} - b_a) - c_{Jaeger,a})^2}. \quad (6)$$

The resulting scaling was applied to the VE data, allowing for direct comparison of model outputs. VE protein data is unavailable for Cad and Tll. For the former, we substituted expression data used by Jaeger *et al.* to fit the original model [13]. For the latter, we substituted Tll mRNA data from the VE and scaled it according to **eqn. 6**.

## Optimization

Using the cost function (eqn. 5), we optimized the full 1D and 3D models against scaled VE data using the Optimization Toolbox in MATLAB R2009a [72]. We began with local searches of the GRN weight matrix  $W$  (containing 42 parameters) using the constrained nonlinear minimization function `fmincon()`. We initialized these searches at the best-fit inferred GRN parameter set of the original modeling study and bounded all parameters within the interval  $[-0.2, 0.2]$  [13]. Parameter and cost function tolerances for stopping criteria were set to zero and the search was allowed to progress for 4200 model evaluations (100 evaluations per parameter), resulting in arrival at local minima. In the case of the DV-asymmetric Bcd model ( $M_{Bcd3D}^{3D}$ ), we subsequently included this locally optimal GRN in the initial population of a global search using genetic algorithms (GAs).

We used the GA as implemented in MATLAB. The population of size twenty genomes (parameter sets) was initialized with nineteen randomized parameter sets and the locally-optimized parameter set found for  $M_{Bcd3D}^{3D}$ . Stopping criteria were specified as a maximum of 100 generations or failure to improve cost function values above a tolerance of  $10^{-6}$ . The latter criterion increments a “stall” counter for each generation that fails to improve the score, ending the GA when the counter reaches fifty [72]. This algorithm was used to search the parameter space while fitting the 3D model incorporating DV-asymmetric Bcd ( $M_{Bcd3D-S}^{3D}$ ).

## Supporting Information

**Figure S1 The VirtualEmbryo geometry.** A three-quarters view of the embryonic geometry with anterior (A), posterior (P), dorsal (D) and ventral (V) poles indicated. (TIF)

**Figure S2 Scaled diffusion constants in DV-symmetric Bcd model  $M_{Bcd1D}^{3D}$ .** The model is insensitive to small changes in the diffusion constant. (**A–G**) Lateral view of VE geometry is shown in rows **A–G** (Gt, Hb, Kni, Kr, Tll at  $t=70$  min, Cad at  $t=56$  min); Column 1 displays output from  $M_{Bcd1D}^{3D}$  evaluated with GRN  $P_{VE}^{1D-F}$  and diffusion constants  $D_a$  scaled by 0.1; Column 2 displays output from  $M_{Bcd1D}^{3D}$  evaluated with GRN  $P_{VE}^{1D-F}$  and diffusion constants  $D_a$  scaled by 0.5; Column 3 displays output from  $M_{Bcd1D}^{3D}$  evaluated with GRN  $P_{VE}^{1D-F}$  and diffusion constants  $D_a$  scaled by 1; Column 4 displays output from  $M_{Bcd1D}^{3D}$  evaluated with GRN  $P_{VE}^{1D-F}$  and diffusion constants  $D_a$  scaled by 2; Column 5 displays output from  $M_{Bcd1D}^{3D}$  evaluated with GRN  $P_{VE}^{1D-F}$  and diffusion constants  $D_a$  scaled by 10. (TIF)

**Figure S3 Scaled diffusion constants in DV-asymmetric Bcd model  $M_{Bcd3D-S}^{3D}$ .** The model is insensitive to small changes in the diffusion constant. (**A–G**) Lateral view of VE geometry is shown in rows **A–G** (Gt, Hb, Kni, Kr, Tll at  $t=70$  min, Cad at  $t=56$  min); Column 1 displays output from  $M_{Bcd3D-S}^{3D}$  evaluated with GRN  $P_{Bcd3D-S}^{3D}$  and diffusion constants  $D_a$  scaled by 0.1; Column displays output from  $M_{Bcd3D-S}^{3D}$  evaluated with GRN  $P_{Bcd3D-S}^{3D}$  and diffusion constants  $D_a$  scaled by 0.5; Column 3 displays output from  $M_{Bcd3D-S}^{3D}$  evaluated with GRN  $P_{Bcd3D-S}^{3D}$  and diffusion constants  $D_a$  scaled by 1; Column 4 displays output from  $M_{Bcd3D-S}^{3D}$  evaluated with GRN  $P_{Bcd3D-S}^{3D}$  and diffusion constants  $D_a$  scaled by 2; Column 5 displays output from  $M_{Bcd3D-S}^{3D}$  evaluated with GRN  $P_{Bcd3D-S}^{3D}$  and diffusion constants  $D_a$  scaled by 10. (TIF)

**Movie S1 Animated Cad pattern formation in the DV-symmetric Bcd model.  $M_{Bcd1D}^{3D}$ .** Model output for Cad evaluated at the parameter set  $P_{VE}^{1D-F}$ . The video spans  $t=0$ –

70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S2 Animated Gt pattern formation in the DV-symmetric Bcd model.**  $M_{Bcd1D}^{3D}$ . Model output for Gt evaluated at the parameter set  $P_{VE}^{1D-F}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S3 Animated Hb pattern formation in the DV-symmetric Bcd model.**  $M_{Bcd1D}^{3D}$ . Model output for Hb evaluated at the parameter set  $P_{VE}^{1D-F}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S4 Animated Kni pattern formation in the DV-symmetric Bcd model.**  $M_{Bcd1D}^{3D}$ . Model output for Kni evaluated at the parameter set  $P_{VE}^{1D-F}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S5 Animated Kr pattern formation in the DV-symmetric Bcd model.**  $M_{Bcd1D}^{3D}$ . Model output for Kr evaluated at the parameter set  $P_{VE}^{1D-F}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S6 Animated Tll pattern formation in the DV-symmetric Bcd model.**  $M_{Bcd1D}^{3D}$ . Model output for Tll evaluated at the parameter set  $P_{VE}^{1D-F}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S7 Animated Cad pattern formation in the DV-asymmetric Bcd model.**  $M_{Bcd3D-S}^{3D}$ . Model output for Cad evaluated at the parameter set  $P_{Bcd3D-S}^{3D}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S8 Animated Gt pattern formation in the DV-asymmetric Bcd model.**  $M_{Bcd3D-S}^{3D}$ . Model output for Gt evaluated at the parameter set  $P_{Bcd3D-S}^{3D}$ . The video spans  $t=0$ –

70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S9 Animated Hb pattern formation in the DV-asymmetric Bcd model.**  $M_{Bcd3D-S}^{3D}$ . Model output for Hb evaluated at the parameter set  $P_{Bcd3D-S}^{3D}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S10 Animated Kni pattern formation in the DV-asymmetric Bcd model.**  $M_{Bcd3D-S}^{3D}$ . Model output for Kni evaluated at the parameter set  $P_{Bcd3D-S}^{3D}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S11 Animated Kr pattern formation in the DV-asymmetric Bcd model.**  $M_{Bcd3D-S}^{3D}$ . Model output for Kr evaluated at the parameter set  $P_{Bcd3D-S}^{3D}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**Movie S12 Animated Tll pattern formation in the DV-asymmetric Bcd model.**  $M_{Bcd3D-S}^{3D}$ . Model output for Tll evaluated at the parameter set  $P_{Bcd3D-S}^{3D}$ . The video spans  $t=0$ –70 min. Concentration colormap is identical to Fig. 4 and spans 0–250 concentration units.

(M4V)

**File S1 Mesh coordinates for the VirtualEmbryo.** Matlab-accessible file containing indexed vertex coordinates and relations defining the triangular elements of the mesh in Figure S1.

(MAT)

**Table S1 GRN Parameter Values.**

(DOC)

**Table S2 Non-GRN Parameters (Unoptimized).**

(DOC)

## Author Contributions

Conceived and designed the experiments: DU JH AR MG CF. Performed the experiments: DU JH. Analyzed the data: DU JH AR MG. Contributed reagents/materials/analysis tools: CF. Wrote the paper: DU JH AR MG CF.

## References

- St Johnston D, Nusslein-Volhard C (1992) The origin of pattern and polarity in the *Drosophila* embryo. *Cell*. pp 201–219.
- Jaeger J (2011) The gap gene network. *Cellular and Molecular Life Sciences*. pp 243–274.
- Kopp A, Blackman R, Duncan I (1999) Wingless, decapentaplegic and EGF receptor signaling pathways interact to specify dorso-ventral pattern in the adult abdomen of *Drosophila*. *Development* 126: 3495–3507.
- Umulis D, Serpe M, O'Connor M, Othmer H (2006) Robust, bistable patterning of the dorsal surface of the *Drosophila* embryo. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11613–11618.
- Anderson K (1987) Dorsal Ventral Embryonic Pattern Genes of *Drosophila*. *Trends in Genetics* 3: 91–97.
- Reinartz J, Sharp D (1995) Mechanism of Eve stripe formation. *Mechanisms of Development*. pp 133–158.
- Surkova S, Kosman D, Kozlov K, Manu, Myasnikova E, et al. (2008) Characterization of the *Drosophila* segment determination morphome. *Developmental Biology*. pp 844–862.
- Akam M (1987) The Molecular-Basis for Metameric Pattern in the *Drosophila* Embryo. *Development* 101: 1–22.
- Ingham P, Arias A (1992) Boundaries and fields in early embryos. *Cell* 68: 221–235.
- Counce S (1961) Analysis of Insect Embryogenesis. *Annual Review of Entomology* 6: 295–312.
- Jaeger J, Sharp D, Reinartz J (2007) Known maternal gradients are not sufficient for the establishment of gap domains in *Drosophila melanogaster*. *Mechanisms of Development*. pp 108–128.
- Irish V, Lehmann R, Akam M (1989) The *Drosophila* posterior-group gene Nanos functions by repressing Hunchback activity. *Nature* 338: 646–648.
- Jaeger J, Surkova S, Blagov M, Janssens H, Kosman D, et al. (2004) Dynamic control of positional information in the early *Drosophila* embryo. *Nature*. pp 368–371.
- Jaeger J, Blagov M, Kosman D, Kozlov K, Manu, et al. (2004) Dynamical analysis of regulatory interactions in the gap gene system of *Drosophila melanogaster*. *Genetics*. pp 1721–1737.
- Small S, Kraut R, Hoey T, Warrior R, Levine M (1991) Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes & Development*. pp 827–839.
- Driever W, Ma J, Nusslein-Volhard C, Ptashne M (1989) Rescue of Bicoid mutant *Drosophila* embryos by Bicoid fusion proteins containing heterologous activating sequences. *Nature*. pp 149–154.
- Driever W, Nusslein-Volhard C (1989) The Bicoid protein is a positive regulator of Hunchback transcription in the early *Drosophila* embryo. *Nature*. pp 138–143.



18. Driever W, Nusslein-Volhard C (1988) The Bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell*. pp 95–104.
19. Lehmann R, Nusslein-Volhard C (1987) Hunchback, a gene required for segmentation of the anterior and posterior region of the *Drosophila* embryo. *Developmental Biology* 119: 402–417.
20. Jurgens G, Wieschaus E, Nusslein-Volhard C, Kluding H (1984) Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* 2 Zygotic loci on the 3rd chromosome. *Wilhelm Roux Archives of Developmental Biology* 193: 283–295.
21. Nusslein-Volhard C, Wieschaus E, Kluding H (1984) Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* 1 Zygotic loci on the 2nd chromosome. *Wilhelm Roux Archives of Developmental Biology* 193: 267–282.
22. Wieschaus E, Nusslein-Volhard C, Jurgens G (1984) Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* 3 Zygotic loci on the X-chromosome and the 4th chromosome. *Wilhelm Roux Archives of Developmental Biology* 193: 296–307.
23. Wieschaus E, Nusslein-Volhard C, Kluding H (1984) Kruppel, a gene whose activity is required early in the zygotic genome for normal embryonic segmentation. *Developmental Biology* 104: 172–186.
24. Nusslein-Volhard C, Wieschaus E (1980) Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287: 795–801.
25. Jackle H, Tautz D, Schuh R, Seifert E, Lehmann R (1986) Cross-regulatory interactions among the gap genes of *Drosophila*. *Nature* 324: 668–670.
26. Nusslein-Volhard C, Frohnhofer H, Lehmann R (1987) Determination of the anteroposterior polarity in *Drosophila*. *Science* 238: 1675–1681.
27. Harding K, Levine M (1988) Gap genes define the limits of Antennapedia and Bithorax gene-expression during early development in *Drosophila*. *Embo Journal* 7: 205–214.
28. Tautz D (1988) Regulation of the *Drosophila* segmentation gene Hunchback by 2 maternal morphogenetic centers. *Nature* 332: 281–284.
29. Hulskamp M, Pfeifle C, Tautz D (1990) A morphogenetic gradient of Hunchback protein organizes the expression of the gap genes Kruppel and Knirps in the early *Drosophila* embryo. *Nature* 346: 577–580.
30. Bronner G, Jackle H (1991) Control and function of terminal gap gene activity in the posterior pole region of the *Drosophila* embryo. *Mechanisms of Development*. pp 205–211.
31. Eldon E, Pirrotta V (1991) Interactions of the *Drosophila* gap gene Giant with maternal and zygotic pattern-forming genes. *Development* 111: 367.
32. Kraut R, Levine M (1991) Mutually repressive interactions between the gap genes Giant and Kruppel define the middle body regions of the *Drosophila* embryo. *Development* 111: 611–621.
33. Capovilla M, Eldon E, Pirrotta V (1992) The giant gene of *Drosophila* encodes a B-zip DNA-binding protein that regulates the expression of other segmentation gap genes. *Development* 114: 99–112.
34. Struhl G, Johnston P, Lawrence P (1992) Control of *Drosophila* body pattern by the Hunchback morphogen gradient. *Cell* 69: 237–249.
35. Hulskamp M, Lukowitz W, Beermann A, Glaser G, Tautz D (1994) Differential regulation of target genes by different alleles of the segmentation gene Hunchback in *Drosophila*. *Genetics* 138: 125–134.
36. Simpsonbrose M, Treisman J, Desplan C (1994) Synergy between the Hunchback and Bicoid morphogens is required for anterior patterning in *Drosophila*. *Cell* 78: 855–865.
37. Ashyraliyev M, Siggins K, Janssens H, Blom J, Akam M, et al. (2009) Gene Circuit Analysis of the Terminal Gap Gene huckebein. *Plos Computational Biology* 5(10): e1000548.
38. Manu, Surkova S, Spirov A, Gursky V, Janssens H, et al. (2009) Canalization of Gene Expression and Domain Shifts in the *Drosophila* Blastoderm by Dynamical Attractors. *Plos Computational Biology* 7(3): e1000049.
39. Manu, Surkova S, Spirov A, Gursky V, Janssens H, et al. (2009) Canalization of Gene Expression in the *Drosophila* Blastoderm by Gap Gene Cross Regulation. *Plos Biology*. pp 591–603.
40. Alves F, Dilao R (2006) Modeling segmental patterning in *Drosophila*: Maternal and gap genes. *Journal of Theoretical Biology* 241: 342–359.
41. Sanchez L, Thieffry D (2001) A logical analysis of the *Drosophila* gap-gene system. *Journal of Theoretical Biology* 212(1): 115–141.
42. Perkins T, Jaeger J, Reinitz J, Glass L (2006) Reverse engineering the gap gene network of *Drosophila melanogaster*. *Plos Computational Biology*. pp 417–428.
43. Ashyraliyev M, Jaeger J, Blom J (2008) Parameter estimation and determinability analysis applied to *Drosophila* gap gene circuits. *BMC Systems Biology* 2: 83.
44. Fomekong-Nanfack Y, Postma M, Kaandorp J (2009) Inferring *Drosophila* gap gene regulatory network: a parameter sensitivity and perturbation analysis. *BMC Systems Biology* 3: 94.
45. Fomekong-Nanfack Y, Kaandorp J, Blom J (2007) Efficient parameter estimation for spatio-temporal models of pattern formation: case study of *Drosophila melanogaster*. *Bioinformatics* 23: 3356–3363.
46. Jaeger J, Reinitz J (2006) On the dynamic nature of positional information. *Bioessays*. pp 1102–1111.
47. Keranen S, Fowlkes C, Luengo Hendriks C, Sudar D, Knowles D, et al. (2006) Three-dimensional morphology and gene expression in the *Drosophila* blastoderm at cellular resolution II: dynamics. *Genome Biology* 7: R124.
48. Luengo Hendriks C, Keranen S, Fowlkes C, Simirenko L, Weber G, et al. (2006) Three-dimensional morphology and gene expression in the *Drosophila* blastoderm at cellular resolution I: data acquisition pipeline. *Genome Biology* 7: R123.
49. Fowlkes C, Luengo Hendriks C, Keranen S, Weber G, Rubel O, et al. (2008) A quantitative spatiotemporal atlas of gene expression in the *Drosophila* blastoderm. *Cell*. pp 364–374.
50. Vakulenko S, Manu, Reinitz J, Radulescu O (2009) Size Regulation in the Segmentation of *Drosophila*: Interacting Interfaces between Localized Domains of Gene Expression Ensure Robust Spatial Patterning. *Physical Review Letters* 103: 168102–5.
51. Gursky V, Jaeger J, Kozlov K, Reinitz J, Samsonov A (2004) Pattern formation and nuclear divisions are uncoupled in *Drosophila* segmentation: comparison of spatially discrete and continuous models. *Physica D-Nonlinear Phenomena*. pp 286–302.
52. Nusslein-Volhard C (1991) Determination of the embryonic axes of *Drosophila*. *Development* 113: 1–10.
53. Carroll S, Winslow G, Twombly V, Scott M (1987) Genes that control dorsoventral polarity affect gene-expression along the anteroposterior axis of the *Drosophila* embryo. *Development* 99: 327–332.
54. Okabe-Oho Y, Murakami H, Oho S, Sasai M (2009) Stable, Precise, and Reproducible Patterning of Bicoid and Hunchback Molecules in the Early *Drosophila* Embryo. *Plos Computational Biology* 5(8): e1000486.
55. Deng JY, Wang W, Lu LJ, Ma J (2010) A Two-Dimensional Simulation Model of the Bicoid Gradient in *Drosophila*. *Plos One* 5.
56. He F, Saunders T, Wen Y, Cheung D, Jiao R, et al. (2010) Shaping a Morphogen Gradient for Positional Precision. *Biophysical Journal*. pp 697–707.
57. Jaeger J (2010) A matter of timing and precision. *Molecular Systems Biology* 6: 427.
58. Jaeger J (2009) Modelling the *Drosophila* embryo. *Molecular Biosystems* 5: 1549–1568.
59. Bieler J, Pozzorini C, Naef F (2011) Whole-embryo modeling of early segmentation in *Drosophila* identifies robust and fragile expression domains. *Biophysical Journal* 101: 287–296.
60. Munteanu A, Sole R (2008) Neutrality and Robustness in Evo-Devo: Emergence of Lateral Inhibition. *Plos Computational Biology* 4(11): e1000226.
61. Wagner A (2008) Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B-Biological Sciences*. pp 91–100.
62. Wagner A (2005) Distributed robustness versus redundancy as causes of mutational robustness. *Bioessays*. pp 176–188.
63. Gregor T, Wieschaus E, McGregor A, Bialek W, Tank D (2007) Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell*. pp 141–152.
64. Gregor T, Bialek W, van Steveninck R, Tank D, Wieschaus E (2005) Diffusion and scaling during early embryonic pattern formation. *Proceedings of the National Academy of Sciences of the United States of America*. pp 18403–18407.
65. Kanodia J, Rikhy R, Kim Y, Lund V, DeLotto R, et al. (2009) Dynamics of the Dorsal morphogen gradient. *Proceedings of the National Academy of Sciences of the United States of America*. pp 21707–21712.
66. Liberman L, Reeves G, Stathopoulos A (2009) Quantitative imaging of the Dorsal nuclear gradient reveals limitations to threshold-dependent patterning in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. pp 22317–22322.
67. Wang C, Lehmann R (1991) Nanos is the localized posterior determinant in *Drosophila*. *Cell*. pp 637–647.
68. Umulis D, Shimmi O, O'Connor M, Othmer H (2010) Organism-Scale Modeling of Early *Drosophila* Patterning via Bone Morphogenetic Proteins. *Developmental Cell* 18: 260–274.
69. COMSOL Multiphysics® 3.5a. COMSOL AB, Palo Alto, CA, USA.
70. Sample C, Shvartsman S (2010) Multiscale modeling of diffusion in the early *Drosophila* embryo. *Proceedings of the National Academy of Sciences of the United States of America* 107: 10092–10096.
71. Grimm O, Wieschaus E (2010) The Bicoid gradient is shaped independently of nuclei. *Development* 137: 2857–2862.
72. MATLAB® R2009a. The Mathworks, Natick, MA, USA.